IL NUOVO CIMENTO **42 C** (2019) 197
DOI 10.1393/ncc/i2019-19197-3

Colloquia: IFAE 2018

Simulating the LHCb hadron calorimeter with generative adversarial networks

D. Lancierini, P. Owen and N. Serra

Physik-Institut, Universität Zürich - Zürich, Switzerland

received 31 January 2019

Summary. — Generative adversarial networks are known as a tool for fast simulation of data. Our aim is to research and develop a physical application of these tools by simulating LHCb hadron calorimeter (HCAL) in order to speed up the Monte Carlo datasets production.

1. – Lepton flavour universality tests

1˙1. *Current status for $R(D^{(*)})$.* – Some of the most interesting phenomena reported by particle physics experiments in the last few years are the numerous hints of lepton flavour universality (LFU) violation observed in semi-leptonic B decays. LFU can be tested with B meson semi-leptonic decays by measuring the ratio of the relative branching fractions involving different lepton flavour types. An example of such a ratio is known as $\mathrm{R}(\mathrm{D}^{(*)})$:

$$\mathrm{R}(\mathrm{D}^{(*)}) = \frac{\mathcal{B}(\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu)}{\mathcal{B}(\mathrm{B} \to \mathrm{D}^{(*)}\mu\nu)} \tag{1}$$

Recent measurements [1] have shown an enhancement of the ratio defined in eq. (1) compared to the Standard Model (SM) prediction [2]. Currently the world average of the combined measurements of $\mathrm{R}(\mathrm{D}^{(*)})$ deviates from the SM prediction by 3.62 σ [3].

This might hint to a LFU-violating contribution of New Physics (NP) origin that couples mainly to the third generation of quarks and leptons, enhancing the $\mathrm{b} \to \tau$ transition.

1˙2. *The $R(D)$ measurement at LHCb.* – The measurement of R(D) is currently ongoing at LHCb and the data collected by the experiment is more than enough to perform the analysis; a more pressing issue is the availability of Monte Carlo (MC) simulated data to describe the effect of the detector on the signal and background processes.

2. – Fast simulation techniques

2·1. *The hadron calorimeter.* – The LHCb hadron calorimeter (HCAL) [4] implements ∼70% of Level 0 trigger in the presence of high-p_t hadrons, and consists of a transversely segmented plastic scintillator composed by cells of two different sizes. It is characterised by moderate energy resolution although sufficient for trigger purposes. The simulation of the LHCb calorimeter response occupies a large fraction of the computing time for MC production. This work addresses the issue of simulating the LHCb calorimeter by parametrising its response using fast machine learning simulation techniques.

2·2. *Adversarial networks.* – The neural network architectures which this work is based on are known as Deep Convolutional Generative Adversarial Networks (DCGANs) and were firstly introduced by Ian Goodfellow in ref. [5]. These architectures, which have been the subject of intense study (see, for example, ref. [6]), are a class of unsupervised neural networks which consist of two neural networks: a generator (G) and a discriminator (D). In the training cycle, G learns a function that maps a uniformly distributed variable z to the dataset x and produces fake samples of data, while D learns to distinguish the fake samples $G(z)$ from the real dataset. These two networks play a non-cooperative minimax game with the following value function $V(G, D)$ [5]:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x\sim p_{data}(x)}\log D(x) + \mathbb{E}_{z\sim p_z(z)}\log(1 - D(G(z))). \tag{2}$$

The training cycle ends when D has a discrimination rate of 50%, and G is trained to reproduce fake samples distributed as close as possible to the real dataset. The existence of such equilibrium is not guaranteed but several attempts have been made to help optimise its search, as shown in ref. [6].

2·3. *Generative models for the LHCb HCAL.* – In order to parametrise the LHCb HCAL using neural networks we tested different architectures on the same dataset. This dataset was generated using PGun and consists of several events in which a single pion hits the calorimeter at different positions and with variable p_t. In order to fully exploit the advantages of a convolutional network, each event in the dataset is converted to an image. The values of non-zero pixels and their locations are determined by the reconstructed energy and location of each calorimeter cell. The architectures tested in this work are Conditional DGCANs [7] and BicycleGANs [8], since they both provide a control on the input as well as a stochastic output.

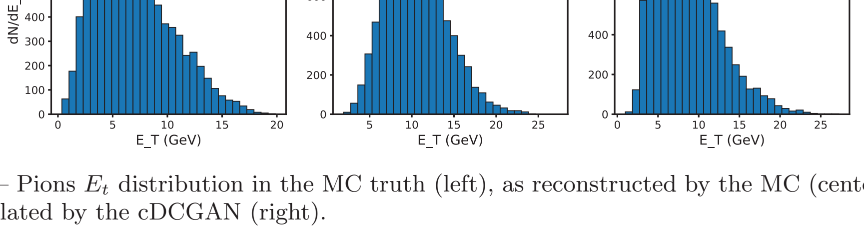

Fig. 1. – Pions E_t distribution in the MC truth (left), as reconstructed by the MC (center) and as simulated by the cDCGAN (right).

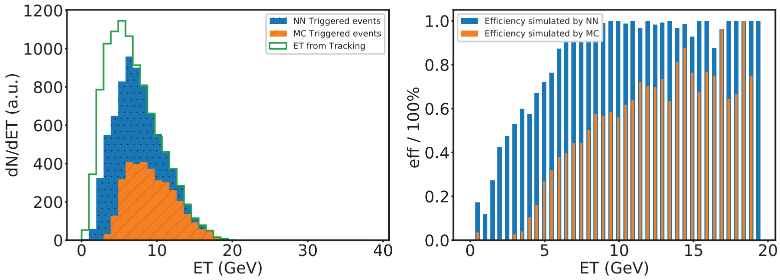

Fig. 2. – Absolute trigger efficiency (left) and relative trigger efficiency (right) comparison between MC and cDCGAN.

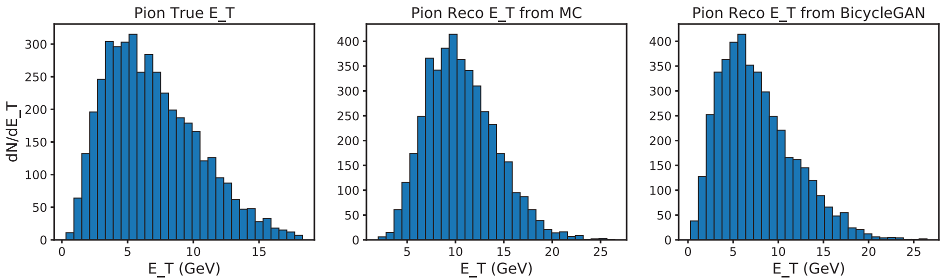

Fig. 3. – Pions E_t distribution in the MC truth (left), as reconstructed by the MC (center) and as simulated by the BicycleGAN (right).

3. – Architectures' comparison

3·1. *Conditional DCGAN (cDCGAN).* – We trained a cDCGAN network on a sample of $\sim 3.6 \times 10^4$ single-pion events with energies spanning $0 \lesssim E_t \lesssim 25\,\text{GeV}$. In the case of the cDCGAN, the network input consists of a string of (x, y, E_t). Once the network training was completed, the network was tested on a test set which is disjoint from the training set. The results are plotted in fig. 1. In each of the panels the distribution of the sum over all energies recorded in each cell per event is plotted. The panel on the left shows the distribution of the true E_t as simulated by the tracking, while the central panel displays the distribution of the HCAL reconstructed E_t as simulated by the MC. The plot on the right shows the reconstructed E_t as simulated by the cDCGAN.

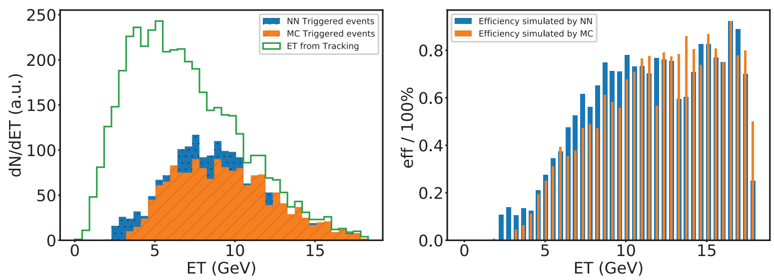

Fig. 4. – Absolute trigger efficiency (left) and relative trigger efficiency (right) comparison between MC and BicycleGAN.

More importantly our aim is to emulate the HCAL trigger efficiency, *i.e.*, the ratio between number of events that fire the trigger and the total number of events. The results for such comparison is shown in fig. 2. The network took and average of 2.98×10^{-3} s to generate one event on an NVidia Tesla P100 GPU.

3˙2. *BicycleGAN*. – We trained a BicycleGAN [8] on the same dataset. This architecture is more complex than the cDCGAN, as it takes as input images rather than strings. The comparison between the distributions of total energy per event is shown in fig. 3.

As shown in fig. 4 the emulation of the trigger efficiency is more accurate in the case of BycycleGAN, although the augmented complexity of the network slows the speed of production to an average of 4.79×10^{-3} s per event on the same hardware.

4. – Conclusions

As shown in this work, generative adversarial networks provide a useful tool for fast MC production, although further research is needed in order to assess the scalability of their application to the simulation of LHCb HCAL response to a full event.

REFERENCES

[1] LHCb Collaboration, *Phys. Rev. Lett.*, **120** (2018) 171802.
[2] Bigi D., Gambino P. and Schacht S., *Phys. Rev. D*, **94** (2016) 094008.
[3] Amhis Y. *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, arXiv:1412.7515; updated results and plots available at `http://www.slac.stanford.edu/xorg/hfag/`.
[4] Pereima D., *J. Instrum.*, **12** (2017) C06016.
[5] Goodfellow I. J. *et al.*, *Generative Adversarial Networks*, arXiv:1406.2661 (2014).
[6] Salimans T. *et al.*, *Improved Techniques for Training GANs*, arXiv:1606.03498 (2016).
[7] Mirza M. *et al.*, *Conditional Generative Adversarial Nets*, arXiv:1411.1784v1 (2014).
[8] Zhu J. *et al.*, *Toward Multimodal Image-to-Image Translations*, arXiv:1711.11586 (2014).